IBM FlashSystem in OLAP Database Environments

IBM Redbooks Solution Guide

IBM® FlashSystem storage systems deliver high performance and efficiency in an easy to integrate offering so that businesses can more readily compete in the market. IBM's recent acquisition of Texas Memory Systems (TMS), an IBM Company, extends IBM's leadership in flash optimized storage solutions. By using patented Variable Stripe RAID™ technology and 2D Flash RAID™, FlashSystem storage systems enhance system resiliency without sacrificing performance or usable capacity.

Figure 1 illustrates the value of the IBM FlashSystem storage infrastructure.
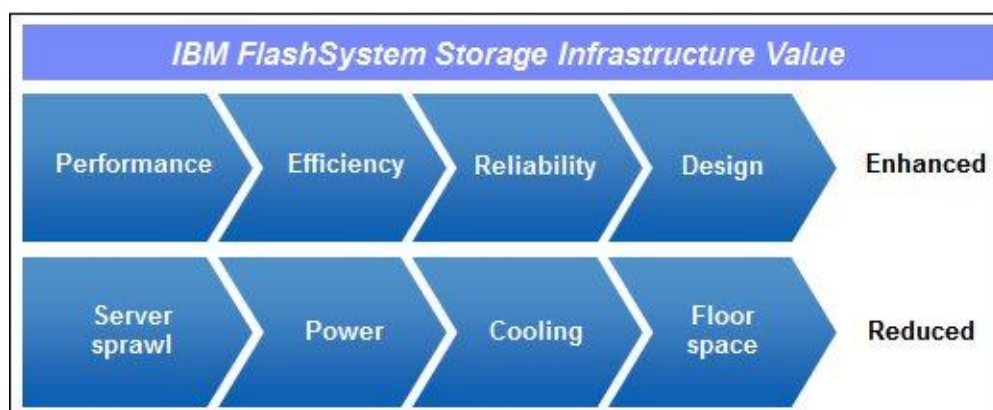

Figure 1. IBM FlashSystem storage infrastructure value

FlashSystem storage systems transform the data center environment and enhance performance and resource consolidation to gain the most from business processes and critical applications. Examples of such processes and applications include online transaction processing (OLTP), business intelligence (BI), online analytical processing (OLAP), virtual desktop infrastructures, high-performance computing, and content delivery solutions (such as cloud storage and video on demand). This guide focuses on IBM flash storage solutions for OLAP database environments, which are based on FlashSystem storage systems.

Did you know?

FlashSystem storage systems deliver over 500,000 read IOPS with less than 100 microseconds of latency and provide up to 20 TB of usable data storage just in 1U of rack space. FlashSystem storage systems also offer enterprise-level availability and reliability with no single point of failure, multiple layers of data correction, chip redundancy, and redundant hot swap components.

Business value

Data warehouses are commonly used with OLAP workloads in decision support systems, such as financial analysis. Unlike OLTP, where transactions are typically relatively simple and deal with small amounts of data, OLAP queries are more complex and process larger volumes of data.

OLAP databases are normally separated from OLTP databases and tend to consolidate historical and reference information from multiple sources. Queries are submitted to OLAP databases to analyze consolidated data from different points of view to make better business decisions in a timely manner.

For OLAP workloads, a fast response time is critical to ensure that strategic business decisions can be made quickly in dynamic market conditions. Delays can significantly increase business and financial risks. Usually, decision making is stalled or delayed because of a lack of accurate, real-time operational data for analytics, which means missed opportunities for the following reasons:

- Inability to gain insight into a business
- Inability to predict business outcomes
- Explosion of volume, variety, and velocity of information

FlashSystem storage systems can help to make businesses more agile and analytics-driven by providing up-to-the-minute analytics based on real-time data, and not yesterday's news.

The delays come primarily from batch data loads and performance issues due to handling heavy complex queries that use I/O resources. A common performance bottleneck in OLAP environments is the I/O that is required for reading massive amounts of data (frequently referred to as *big data*) from storage for processing in the OLAP database server. The server ability to process this data is usually a nonfactor because they typically have significant amounts of RAM and processing power, parallelizing tasks across the computing resources of the servers.

In general, clients might experience the following challenges in OLAP environments:

- Slow query execution and response times, which delay business decision making
- Dramatic growth in data, which requires deeper analysis

FlashSystem storage systems can help to address these challenges in the following ways:

- Dramatically boosting the performance of OLAP workloads with distributed scale-out architecture, providing almost linear and virtually unlimited performance and capacity scalability

- Significantly improving response time for better and timely decision making

Solution overview

An OLAP solution with FlashSystem storage systems consists of the following components:

- Database servers (IBM System x® or IBM Power Systems™) to run data management software such as IBM DB2®, Microsoft SQL Server, or Oracle Database

- Flash storage systems (IBM FlashSystem 720 or IBM FlashSystem 820) to host the entire data set or partitioned subsets of data

- A private network (such as 10 Gb Ethernet or QDR/FDR InfiniBand) that is used to provide high-speed connectivity across database servers in a cluster

- A storage area network (SAN) that is used to provide connectivity across database servers and storage systems

IBM DB2 for Linux, UNIX, and Windows is the database of choice for robust, enterprise-wide solutions that handle high-volume workloads. It is optimized to deliver industry-leading performance, while lowering costs. IBM servers that run DB2 are proven performance leaders. DB2 uses and optimize multiple threads automatically, with no change to applications. The unique clustering design of DB2 provides near-linear scalability, continuous availability, and simplified management.

IBM System x servers are developed, tested for quality, and certified by IBM, and they are backed by IBM's incomparable worldwide service and support. IBM System x servers deliver business value over the long term because of advancements in scalability, reliability, and performance, which are combined with flexible configuration options, energy efficient components, and robust systems management tools.

IBM eX5 servers, such as IBM System x3850 X5, are designed for mission-critical enterprise-class workloads such as databases. These systems use open, industry-standard components, which are combined with IBM eX5 technologies, to provide leadership performance, scalability, and reliability.

Ideally suited for compute-intensive workloads, IBM Power Systems deliver leadership performance and scalability in its class. An integrated approach to the design, development, and testing of each IBM POWER® server, blade, or compute node ensures the resiliency that is required for today's IT infrastructure. All Power Systems server models include innovative reliability, availability, and serviceability features that help you avoid unplanned downtime. And, with capacity on demand, hot-node add, and IBM Active Memory™ Expansion, Power Systems enterprise servers ensure that you can keep your most important applications available, even as you add capacity to handle new business demands.

FlashSystem offerings that are used in OLAP environments include FlashSystem 720 and FlashSystem 820 purpose-built external flash storage systems. These storage systems feature one of the industry's most dense GB capacity, bandwidth, and IOPS per 1U rack space ratios for a protected storage, in addition to low power consumption.

Figure 2 shows FlashSystem 720, high availability SLC Flash system.

Figure 2. FlashSystem 720

Figure 3 shows FlashSystem 820, high availability eMLC Flash system.

Figure 3. FlashSystem 820

FlashSystem 720 and FlashSystem 820 have the following common features:

- 1U form factor
- Two dual-port 8 Gbps Fibre Channel controllers or dual-port QDR InfiniBand host interface modules
- Patented Variable Stripe RAID
- 2D Flash RAID: RAID-5 across flash chips and RAID-5 across flash modules
- Hot swappable flash
- Redundancy for power, data, and management
- No single point of failure

FlashSystem 720 has the following features:

- Enterprise-grade SLC flash
- 5 or 10 TB of usable storage with system-level RAID and spare flash modules
- 525,000 IOPS (4 KB reads) and 400,000 IOPS (4-KB writes)
- 25 µs (writes) and 100 µs (reads) latency
- 5 GBps bandwidth (reads) and 4 GBps (writes)
- Low power (350 watts)

FlashSystem 820 has the following features:

- Enterprise-grade eMLC flash
- 10 or 20 TB of usable storage with system-level RAID and spare flash modules
- 525,000 IOPS (4 KB reads) and 280,000 IOPS (4 KB writes)
- 25 µs (writes) and 110 µs (reads) latency
- 5 GBps bandwidth (reads) and 2.8 GBps (writes)
- Low power (300 watts)

Solution architecture

IBM flash storage solutions for OLAP use a distributed server and storage scale-out approach. This approach satisfies high bandwidth requirements and provides unlimited performance and capacity growth capabilities, matching the growing volumes of data that is being processed.

FlashSystem storage systems are connected to the database hosting platforms by using Fibre Channel SAN. Server hosts, or nodes, are interconnected with the isolated high-speed network (such as 10 Gb Ethernet with IBM RackSwitch™ G8124E) that is used for the inter-node data exchange. Each node runs a copy of the OLAP database application, and the analyzed data set is partitioned and distributed across the storage systems. Depending on the database management software that is used and its architecture, each node might have access to only a certain portion of data, or all nodes can have access to all data that is stored on the external storage. OLAP queries are distributed across nodes and processed in parallel. Figure 4 illustrates this architecture.
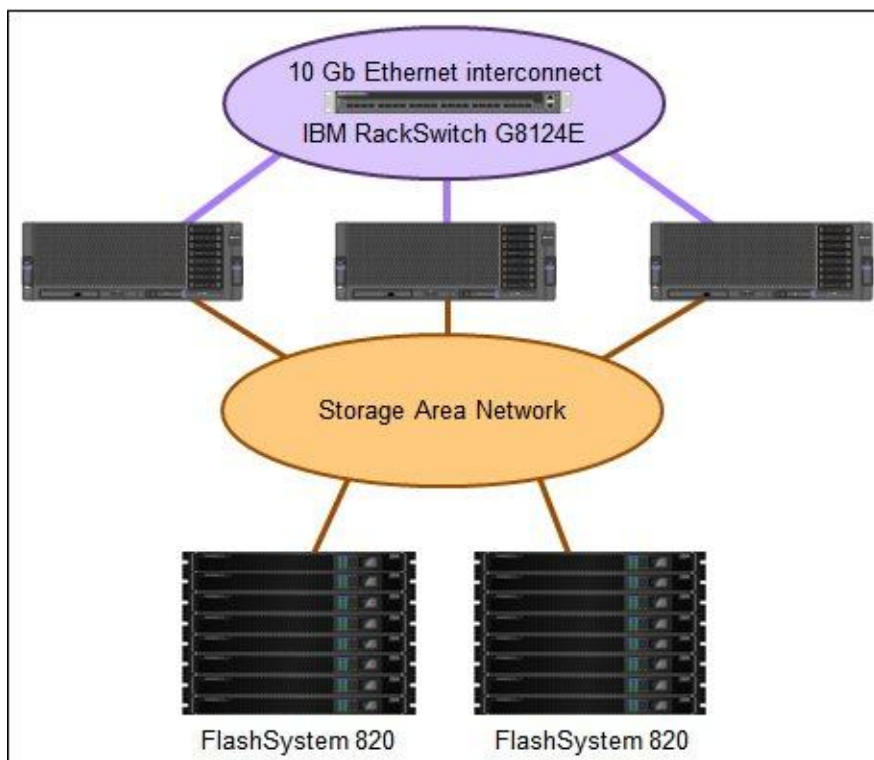

Figure 4. Scale-out OLAP architecture

This solution can scale seamlessly by adding more FlashSystem storage systems and nodes. In such a case, storage capacity and I/O bandwidth are incremented linearly with the increasing number of storage devices, which can help to eliminate storage I/O bottlenecks in OLAP workloads.

Usage scenarios

OLAP applications can be used for risk assessment, business intelligence and reporting, exploration and visualization, predictive analytics, and other similarly profiled industry and functional applications. For example, consider a data warehouse solution that can store and process 160 TB of data. With eight FlashSystem 820 storage systems, the consolidated FlashSystem storage infrastructure can store up to 160 TB of a protected data with up to 26.4 GBps storage I/O aggregated bandwidth.

Integration

An excellent example of a scale-out approach in the OLAP environment is IBM InfoSphere Warehouse. InfoSphere Warehouse is powered by the DB2 for Linux, UNIX, and Windows data server. With its massively scalable, shared-nothing architecture, DB2 provides high performance for mixed-workload query processing of relational and basic XML data. Such advanced features as database and table partitioning, compression, multidimensional clustering (MDC), materialized query tables (MQT), and OLAP capabilities make DB2 a powerful engine for operational warehousing.

InfoSphere Warehouse provides advanced capabilities for database partitioning, so that IT users have multiple ways to distribute data across servers for large-scale parallelism and linear scalability. The shared-nothing architecture of DB2 helps ensure that performance will not degrade as the warehouse grows. Also, because InfoSphere Warehouse can physically cluster data on multiple dimensions, order data by value range, and limit I/O to relevant data partitions, it helps reduce the work that is needed to resolve many queries.

The architecture of the InfoSphere Warehouse database server solution with FlashSystem 820 is shown in the following figure.
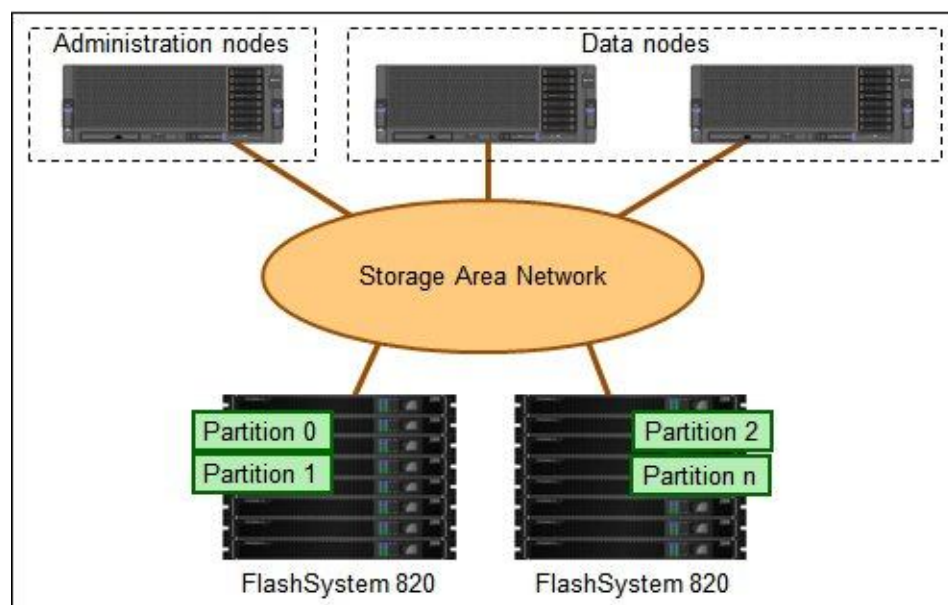

Figure 5. InfoSphere Warehouse database server solution with FlashSystem 820

InfoSphere Warehouse transparently splits the database across multiple partitions stored on FlashSystem logical volumes, and it uses computing power of multiple servers to satisfy requests for large amounts of information. SQL statements are automatically decomposed into subrequests that are run in parallel across each database partition. Results of the subrequests are joined to provide final results.

Supported platforms

FlashSystem storage systems support a wide range of operating systems (Windows Server 2003 and 2008, Linux, and IBM AIX®), hardware platforms (System x, Power Systems, and x86 servers not from IBM), HBAs, and SAN fabrics. For specific information, see the System Storage Interoperation Center (SSIC): http://ibm.com/systems/support/storage/ssic

Ordering information

For FlashSystem 720 and FlashSystem 820 ordering information, see the following IBM Redbooks®
Product Guide:

- *IBM FlashSystem 720 and IBM FlashSystem 820*
  http://www.redbooks.ibm.com/abstracts/tips1003.html

Related information

For more information, see the following documents:

- IBM FlashSystem product page
  http://www.ibm.com/storage/flash

- *IBM FlashSystem 720 and IBM FlashSystem 820 Product Guide*, TIPS-1003
  http://www.redbooks.ibm.com/abstracts/tips1003.html

- IBM System x product page
  http://www.ibm.com/systems/x

- IBM Power Systems product page
  http://www.ibm.com/systems/power

- IBM InfoSphere Warehouse Information Center
  http://bit.ly/SC2IWU

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service. IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you. This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you. Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurement may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

This document was created or updated on April 16, 2013.

Send us your comments in one of the following ways:
- Use the online **Contact us** review form found at:
  ibm.com/redbooks
- Send your comments in an e-mail to:
  redbook@us.ibm.com
- Mail your comments to:
  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400 U.S.A.

This document is available online at http://www.ibm.com/redbooks/abstracts/tips0974.html .

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

Active Memory™
AIX®
DB2®
IBM®
POWER®
Power Systems™
RackSwitch™
Redbooks®
Redbooks (logo)®
System x®

The following terms are trademarks of other companies:

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.